

10-2-2014

At the southeast fringe of the Bantu expansion: genetic diversity and phylogenetic relationships to other sub-Saharan tribes

Diane Rowold

Foundation for Applied Molecular Evolution

Ralph Garcia-Bertrand

Colorado College

Silvia Calderon

New York University

Luis Rivera

College of Health Sciences, Florida International University

David Perez Benedico

Miami Dade College

See next page for additional authors

Follow this and additional works at: https://digitalcommons.fiu.edu/all_faculty

Recommended Citation

Rowold, Diane; Garcia-Bertrand, Ralph; Calderon, Silvia; Rivera, Luis; Benedico, David Perez; Alfonso Sanchez, Miguel A.; Chennakrishnaiah, Shilpa; Varela, Mangela; and Herrera, Rene J., "At the southeast fringe of the Bantu expansion: genetic diversity and phylogenetic relationships to other sub-Saharan tribes" (2014). *All Faculty*. 74.
https://digitalcommons.fiu.edu/all_faculty/74

Authors

Diane Rowold, Ralph Garcia-Bertrand, Silvia Calderon, Luis Rivera, David Perez Benedico, Miguel A. Alfonso Sanchez, Shilpa Chennakrishnaiah, Mangela Varela, and Rene J. Herrera



At the southeast fringe of the Bantu expansion: genetic diversity and phylogenetic relationships to other sub-Saharan tribes



Diane Rowold^e, Ralph Garcia-Bertrand^{a,*}, Silvia Calderon^{c,1}, Luis Rivera^{b,1}, David Perez Benedicto^{d,1}, Miguel A. Alfonso Sanchez^f, Shilpa Chennakrishnaiah^g, Mangela Varela^a, Rene J. Herrera^a

^a Biology Department, Colorado College, Colorado Springs, CO 80903, USA

^b College of Health Sciences, Florida International University, Miami, FL 33199, USA

^c College of Dentistry, New York University, New York, NY 10010, USA

^d Department of Biology, Miami Dade College, Miami, FL 33199, USA

^e Foundation for Applied Molecular Evolution, Gainesville, FL 32601, USA

^f Departamento de Genética y Antropología Física, Facultad de Ciencia y Tecnología, Universidad del País Vasco, 48080 Bilbao, Bizkaia, Spain

^g Department of Human Genetics, McGill University, Montreal, Quebec H3Z 2Z3, Canada

ARTICLE INFO

Article history:

Received 14 May 2014

Revised 6 July 2014

Accepted 15 August 2014

Available online 2 October 2014

Keywords:

Bantu

Africa

Phylogenetics

Y-chromosome

Y-STRs

Population genetics

ABSTRACT

Here, we present 12 loci paternal haplotypes (Y-STR profiles) against the backdrop of the Y-SNP marker system of Bantu males from the Maputo Province of Southeast Africa, a region believed to represent the southeastern fringe of the Bantu expansion. Our Maputo Bantu group was analyzed within the context of 27 geographically relevant reference populations in order to ascertain its genetic relationship to other Bantu and non Bantu (Pygmy, Khoisan and Nilotic) sub-equatorial tribes from West and East Africa. This study entails statistical pair wise comparisons and multidimensional scaling based on YSTR Rst distances, network analyses of Bantu (B2a-M150) and Pygmy (B2b-M112) lineages as well as an assessment of Y-SNP distribution patterns. Several notable findings include the following: 1) the Maputo Province Bantu exhibits a relatively close paternal affinity with both east and west Bantu tribes due to high proportion

Abbreviations: SNP, single nucleotide polymorphism; STR, short tandem repeat; mtDNA, mitochondrial DNA; MAP, Maputo Province; PCR, polymerase chain reaction; RFLP, restriction fragment length polymorphism; MJ, medial joining; MDS, multi-dimensional scaling; TMRCA, time of most recent common ancestor

* Corresponding author at: Biology Department, Colorado College, 14 East Cache La Poudre Street, Colorado Springs, CO 80903-3294, USA. Tel.: +1 719 389 6402; fax: +1 719 389 6940.

E-mail address: RBertrand@coloradocollege.edu (R. Garcia-Bertrand).

¹ These authors contributed equally to the manuscript.

<http://dx.doi.org/10.1016/j.mgene.2014.08.003>

2214-5400/© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

of Bantu Y chromosomal markers, 2) only traces of Khoisan (1.3%) and Pygmy (1.3%) markers persist in the Maputo Province Bantu gene pool, 3) the occurrence of R1a1a-M17/M198, a member of the Eurasian R1a-M420 branch in the population of the Maputo Province, may represent back migration events and/or recent admixture events, 4) the shared presence of E1b1b1-M35 in all Tanzanian tribes examined, including Bantu and non-Bantu groups, in conjunction with its nearly complete absence in the West African populations indicate that, in addition to a shared linguistic, cultural and genetic heritage, geography (e.g., east vs. west) may have impacted the paternal landscape of sub-Saharan Africa, 5) the admixture and assimilation processes of Bantu elements were both highly complex and region-specific.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

Introduction

Bantu encompasses a group of related languages belonging to the Niger–Congo family (Greenberg, 1972) with wide distribution throughout sub-Saharan Africa. It is believed that the proto-Bantu language originated in West Africa in what is now North Cameroon about 5000 years ago (ya) (Greenberg, 1972; Vansina, 1995). Approximately 4000 to 3000 ya, the Bantu-speaking people of West Africa initiated a major human Diaspora and associated cultural transformation that rapidly propagated agriculture and iron work along with the Bantu language to most sub-equatorial Africa (Berniell-Lee et al., 2006; Desmond and Brandt, 1984; Diamond, 1997; Newman, 1995; Phillipson, 1993). Today, the term Bantu is associated with a culture as well. It is theorized that the Bantu demographic expansion proceeded in waves along two primary routes (Berniell-Lee et al., 2006). One path of dispersal transpired from the Bantu homeland along a southwestern course, whereas a second migration, also from North Cameroon, followed a southeastern trajectory (Berniell-Lee et al., 2006; Diamond, 1997; Newman, 1995) and reached its fringes in Southeast Africa as recent as 300 ya (Vansina, 1995). However, more recent data argues for an initial single southwestern migration from the Bantu homeland and a subsequent longitudinal dispersion eastward (Alves et al., 2011; De Filippo et al., 2012; Pakendorf et al., 2011; Russell et al., 2014). It has been proposed that limited agricultural land and overpopulation were the primary motivations for this mass migration (Oliver, 2009). During this extensive geographical and cultural diffusion, the Bantu migrants encountered and interacted with the indigenous sub-Saharan tribes practicing animal husbandry or hunting gathering (Cavalli-Sforza, 1986).

The presence of Bantu-specific markers beyond their homeland suggests that in addition to facilitating an expansive acculturation process in which the Bantu technological advances and lifestyle spread across the African continent, the expansion also included geographical dispersal of the Bantu-speaking people (De Filippo et al., 2012). The degree of Bantu-specific signals in the putative areas which were colonized seems to be population and region specific. In return, the Bantu migrants not only assimilated many elements of the indigenous cultures into the Bantu lifestyle but also expanded their gene pool by intermarriage with the native people encountered during the colonization of sub-equatorial Africa (Berniell-Lee et al., 2006; Ehret, 2001; Tishkoff et al., 2007). An interesting facet of this dispersal is the relative level of gene flow among Bantu and non-Bantu populations occurring in the different regions of contact. However, due to the limited and fragmentary nature of the genetic data from Africa in which individual studies report on different populations and marker systems, many of the demographic aspects of the Bantu expansion are highly debated and unresolved. Thus, the question concerning the relative contributions of the Bantu versus those of the indigenous people to the genetic makeup of the extant sub-Saharan African populations remains largely unanswered.

The available mtDNA data indicates that varying degrees of admixture have occurred between the Bantus and native inhabitants, depending on the indigenous tribes involved, location and the marker employed. In several sub-Saharan Bantu-speaking groups, only traces of ancient mtDNA lineages have been detected (Batini et al., 2007; Quintana-Murci et al., 2008) while in others, such as the Khoisan-speaking Southwest

African populations, a complete replacement of the ancestral by Bantu mtDNA haplogroups is observed (Beleza et al., 2005). In contrast, much higher frequencies of non-Bantu mtDNA lineages are reported in the Bantu groups from Gabon and Cameroon (Berniell-Lee et al., 2009) and Pygmy haplogroups, such as L1c1a, are present in Bantu agriculturalists but Bantu mtDNA in Pygmy groups is rare (Batini et al., 2007; Quintana-Murci et al., 2008). Also, ancestral mtDNA signals persist in the Talta and Mijakeuda, two Bantu-speaking tribes from southeastern Kenya (Batai et al., 2013).

Also, Y-specific markers indicate that low levels of ancient paternal lineages persist in a number of extant Bantu-speaking African tribes (Cruciani et al., 2002; De Filippo et al., 2012; Tishkoff et al., 2007; Underhill et al., 2000). For example, traces (1%) of the local hunter-gatherer Pygmy Y-haplogroups (B2b and A) are observed in the Bantu populations of Gabon and Cameroon (Berniell-Lee et al., 2009). However, in the Khoisan-speaking populations of Southwest Africa, there appears to be a total displacement of the ancestral Y chromosomal lineages by Bantu-specific haplogroups (Beleza et al., 2005), a scenario very similar to present day mtDNA distributions in these people, as discussed above. Thus, in general, the genetic heterogeneity existing across the Bantu domain of sub-Saharan Africa suggests that the admixture and assimilation processes were both highly complex and region-specific.

The Southeast African Bantu populations lie at the outermost fringes of the expansion, and consequently, their genetic composition should reflect centuries of gene flow with non-Bantu groups encountered along the journey southward and eastward. It follows that a genetic study of populations in this important region of Africa and their phylogenetic affinities with relevant Bantus and non-Bantu sub-Saharan African groups may provide valuable information on the degree of gene flow and, perhaps, shed some light on the specific route(s) traveled as well.

The Y-SNP and Y-STR marker systems are useful tools to ascertain human ancestry and population dynamics. Y-SNPs provide reliable ancestral signals and allow for the delineation of migratory paths. Y-STRs are particularly useful in the study of recent human evolution due to their hyper-variability, which enable high levels of phylogenetic resolution (Rowold and Herrera, 2003). To date, to our knowledge, only a handful of Y-STR studies (Alves et al., 2003; Carvalho et al., 2010; Gusmão et al., 2003; Pereira et al., 2002; Sanchez-Diz et al., 2003) have examined the genetic diversity of Bantu populations from the Maputo Province in Southeast Africa. Although these articles provide valuable genetic data, the investigations are limited in scope, uniformity of specific markers across the studies and, for the most part, by the low number of Y-STR loci examined, the lack of both Y-SNP and Y-STR data as part of individual studies as well as phylogenetic comparisons to other continental African groups, Bantu and non-Bantus. To fill this void, we report here 12 loci Y-STR profiles of Bantu individuals from the Maputo Province of Southeast Africa (designated from now on in this article as MAP). In addition, we genotyped the samples with the Y-SNP marker system and conducted phylogenetic comparisons to assess the MAP's genetic relationships to 27 ethnologically well characterized, geographically targeted African groups previously typed for the same Y-SNP and Y-STR loci. From this newly acquired genetic information, the relative proportion of Bantu paternal haplotypes to those representing the region's original inhabitants may be determined which, in turn, may enable inferences regarding the Bantu expansion.

Materials and methods

Sample collection and DNA isolation

Buccal swabs were collected with informed consent from 78 unrelated Bantu-speaking males residing in the Maputo Province, a region in Southeast Africa. The genealogical information of each donor was recorded for a minimum of two generations in order to ascertain their regional ancestry and to ensure non-relatedness among donors. The 78 male individuals sampled belong to the following Bantu-speaking tribes: Shangana (33), Ronga (20), Tswa (10), Chuabo (5), Ndau (2), Makonde (1) and Sena (1). The DNA was isolated as previously described (Calderon et al., 2013) and stored at -80°C .

Published data

A total of 27 geographically relevant reference populations were selected for genetic comparison including West African and East African Bantus as well as six non-Bantu groups (Pygmy populations from

Gabon and Cameroon, and the Burunge, Datog, Hadza and Sandawe from Tanzania). The 27 reference populations were typed for the same 12 Y-STR loci and binary Y-SNP markers genotyped for the MAP population in this study. A list of populations, abbreviations, sample number, geographical regions, language affiliations and references is provided in Table 1. A map of Africa indicating the locations of the populations is included in Fig. 1.

Y-STR and Y-SNP genotyping

All samples were genotyped at twelve Y-STR loci (DYS19, *DYS385a/b*, *DYS389I*, *DYS389II*, *DYS390*, *DYS391*, *DYS392*, *DYS393*, *DYS437*, *DYS438*, and *DYS439*) using an AmpFISTR® Yfiler™ kit (Applied Biosystems, Foster City, CA) in accordance with the manufacturer's specified instructions. PCR products were separated by capillary electrophoresis on an ABI Prism 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) and the allelic categories of the separated Y-STR fragments were determined by the GeneMapper® v.3.2 software. In all analyses, the size of the *DYS389I* allele was subtracted from that of the *DYS389II* fragment.

Bi-allelic markers were assessed in a hierarchical order using standard methods including PCR-RFLP (Luis et al., 2004), allele-specific PCR (Martinez et al., 2007; Regueiro et al., 2006), the YAP polymorphic *Alu* insertion (PAI) (Hammer and Horai, 1995) and direct sequencing (Gayden et al., 2007) to assign individual samples to their respective Y-haplogroups. The Y-haplogroups were ascertained after genotyping the following final SNPs: A1b1b2b1-M118, B2a1a-M109, B2b-M112, E1b1a1a1a-M58,

Table 1
List of populations.

Name	Code	N	Location	Language	Reference
Bantu					
West African					
Akele	AKE	50	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Benga	BEN	48	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Duma	DUM	46	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Eshira	ESH	42	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Eviya	EVI	24	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Fang	FAN	60	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Galoa	GAL	47	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Kota	KOT	53	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Makina	MAK	43	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Ndumu	NDU	36	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Ngumba	NGU	24	Cameroon	Bantu	15 Berniell-Lee et al. (2009)
Nzebi	NZE	57	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Obamba	OBA	47	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Orungu	ORU	21	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Punu	PUN	58	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Shake	SHA	43	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Teke	TEK	48	Gabon	Bantu	15 Berniell-Lee et al. (2009)
Tsogo	TSO	60	Gabon	Bantu	15 Berniell-Lee et al. (2009)
East African					
Rwanda	RWA	67	Rwanda	Bantu	44 Balamurugan and Duncan (2012)
Maputo	MAP	78	Southeast Africa	Bantu	Present Study
Sukuma	SUK	30	Tanzania	Bantu	12 Tishkoff et al. (2007)
Turu	TUR	20	Tanzania	Bantu	12 Tishkoff et al. (2007)
Non-Bantu					
Pygmy					
Baka (Pygmy)	BAK	33	Gabon	Ubangian	15 Berniell-Lee et al. (2009)
Bakola (Pygmy)	BKO	22	Cameroon	Bantu	15 Berniell-Lee et al. (2009)
East African					
Burunge	BUR	23	Tanzania	Nilotic	12 Tishkoff et al. (2007)
Datog	DAT	31	Tanzania	Cushitic	12 Tishkoff et al. (2007)
Hadza	HAD	54	Tanzania	Khoisan	12 Tishkoff et al. (2007)
Sandawe	SAN	67	Tanzania	Khoisan	12 Tishkoff et al. (2007)

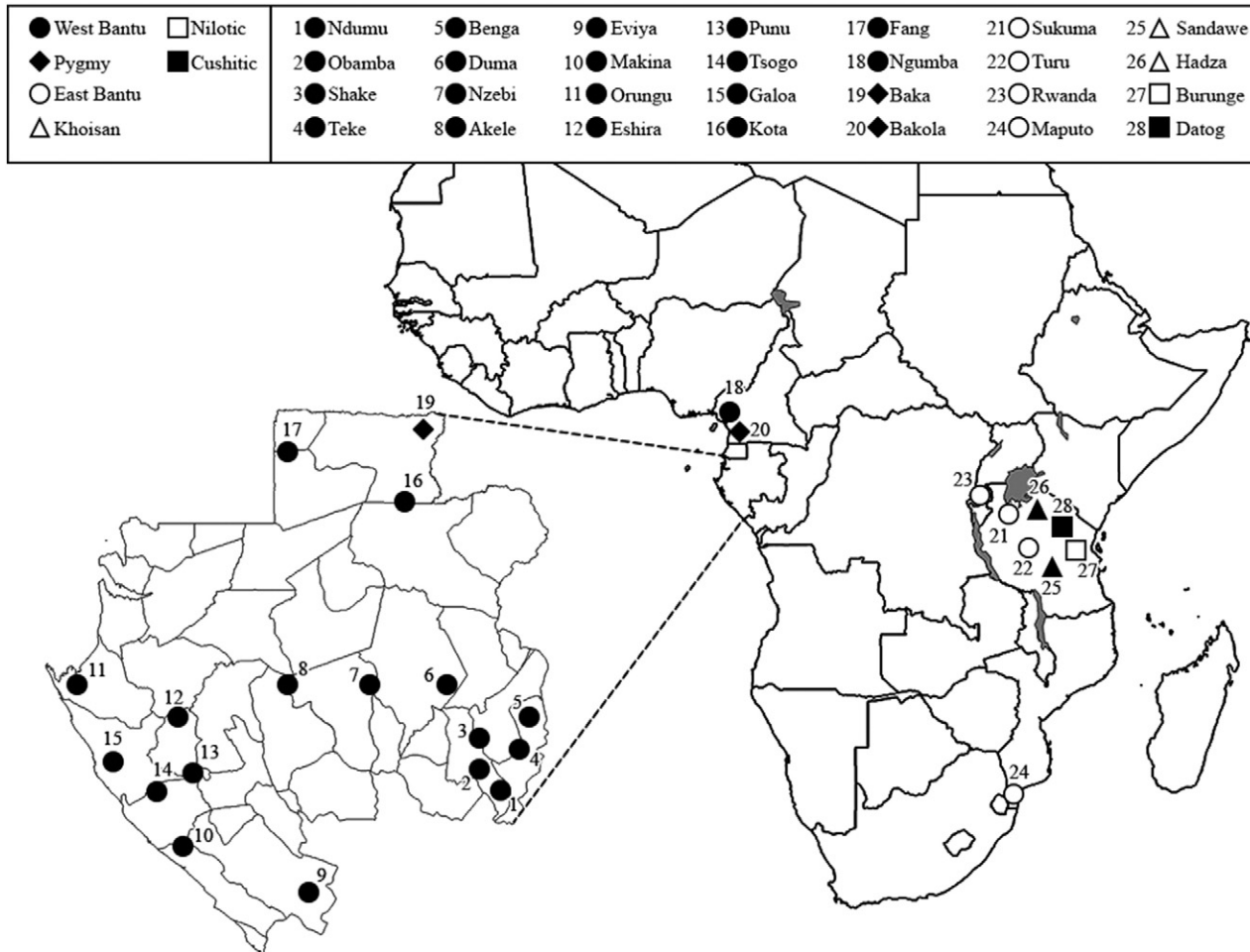


Fig. 1. Map of sub-Sahara Africa indicating geographical location and linguistic groups. Geometric figures represent linguistic affiliations and geographic locations.

E1b1a1a1f1a1-U174, E1b1a1a1g1a-U290, E1b1a1a1g1b-P59, E1b1a1a1g1c-M154, E1b1b2a1-M34, E2b-M98, E2b1-M85, E2b1a-M200, E2b1a1-P45, E2b1a2-P258, and R1a1a-M198. The Y-SNP haplogroup assignment and nomenclature is in accordance with the Y Chromosome Consortium and subsequent revisions (Karafet et al., 2008; Myres et al., 2011; Underhill et al., 2010; Y Chromosome Consortium., 2002).

Statistical and phylogenetic analyses

Pair wise genetic distances (Rst values) and corresponding *P*-values of all populations based on the ten Y-STR loci (excluding DYS385a/b) were generated using the Arlequin v3.5 software package (Excoffier et al., 2005). Pair wise population comparisons were tested at a significance level of 0.005 and 1000 permutations (Kayser et al., 2003) and the Bonferroni correction was applied in order to account for potential type I errors ($\alpha = 0.005/378 = <0.000014$). The DYS385 locus and all samples carrying microvariants were excluded from the Rst calculations. A multidimensional scaling (MDS) plot based on this Rst matrix was constructed using the XLSTAT software from Addinsoft Corp (<http://www.xlstat.com>).

Evolutionary (Zhivotovsky et al., 2004) and genealogical (Goedbloed et al., 2009) mutation rates were employed to generate time estimates based on individuals within the B2a-M150 and B2b-M112 haplogroups (Bantu- and Pygmy-specific markers, respectively). The B2a-M150 and B2b-M112 Media Joining (MJ) networks were generated utilizing the NETWORK 4.6.1.1 software program (www.fluxus-engineering.com) using the twelve loci Y-STR haplotype profiles of the relevant African populations. The R1a1a-M198 Network was done with the same NETWORK 4.6.1.1 program using 17 YSTR loci (the same 12 loci utilized to generate the B2a-M150 and B2b-M112 projections plus DYS456, DYS458, DYS635, GATA-H4 and DY5448). In the network calculations, the Y-STR loci were weighted inversely to the size variance (Martinez et al., 2007; Regueiro et al., 2013). The simplest possible projections were obtained by subjecting the resulting MJ networks to post processing using maximum parsimony parameters.

Results

Haplogroup and haplotype frequencies

The haplogroups and complete haplotypes of the individuals genotyped are provided in Supplementary Table 1. As indicated by the Y-SNP haplogroup frequencies (Table 2), the E1b1a1a1-M180, and B2a1a-M109 mutations, both of which are signatures of the Bantu expansion, rank as the two most abundant Y-SNP haplogroups of the MAP population (at 71.8% and 14.1%, respectively). These haplogroups are followed in

Table 2
Maputo Bantu (MAP) Y-SNP haplogroup frequencies.

Maputo Bantus (N = 78)		
Hg	n	%
A1b1b2b1	1	1.28
B2a1a	11	14.10
B2b	1	1.28
E1b1a1a1a	2	2.56
E1b1a1a1f1a1	17	21.79
E1b1a1a1f1a1c	2	2.56
E1b1a1a1g1	19	24.36
E1b1a1a1g1a	4	5.13
E1b1a1a1g1b	3	3.85
E1b1a1a1g1c	9	11.54
E1b1b2a1	1	1.28
E2b	1	1.28
E2b1	3	3.85
E2b1a	2	2.56
R1a1a	2	2.56

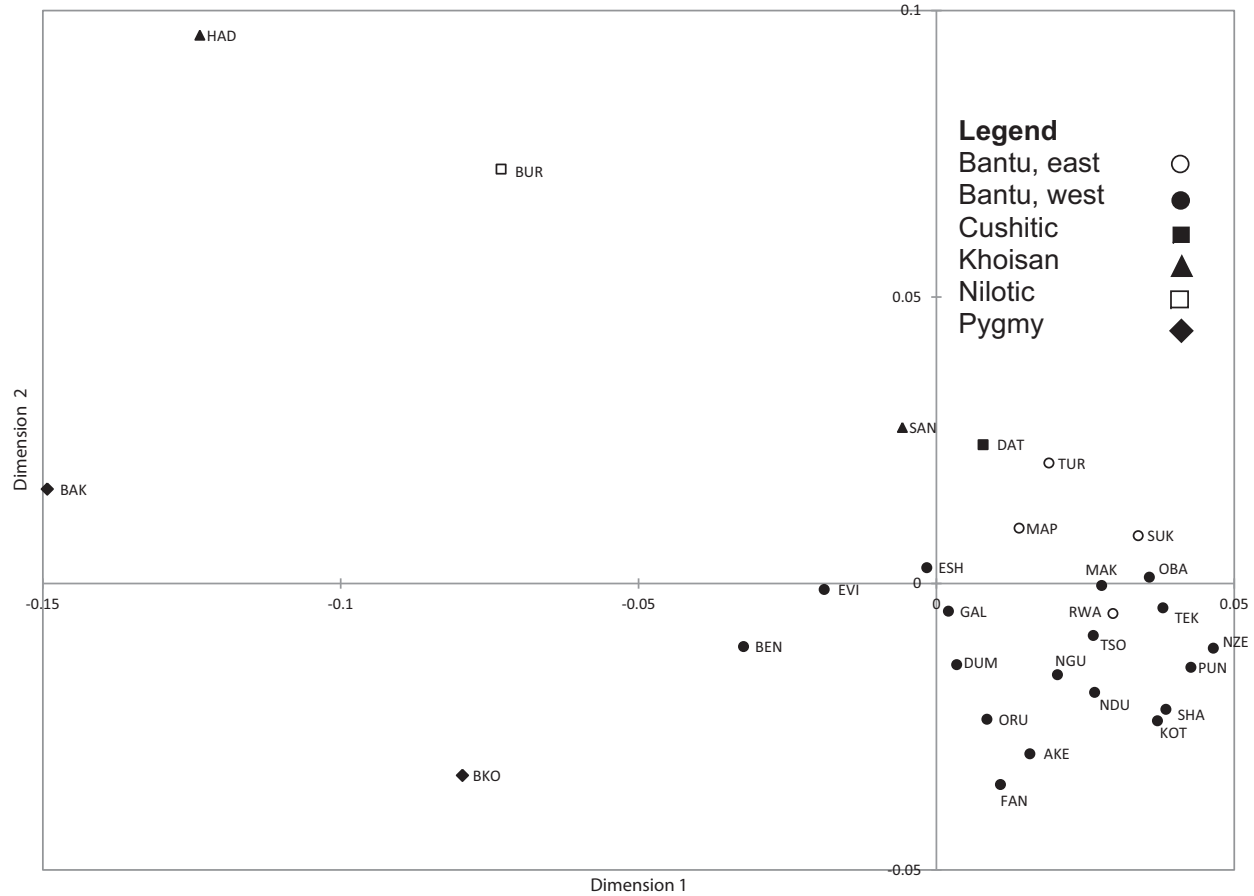


Fig. 2. Multidimensional Scaling plot based on 12 loci Y-STR profiles of African males from 28 populations (see Table 1 for population abbreviations and other information). Configuration (Kruskal's stress (1) = 0.097).

frequency by the Bantu and Eurasian markers E2b (7.7%) that include E2b-M54, E2b1-M85 and E2b1a-M200, and R1a1a-M198 (2.6%), respectively. The ancient East African haplogroup A1b1b2b1-M118 and the Pygmy haplogroup B2b-M112 are represented by one (1.3%) individual each.

Population pair wise Rst distance comparisons

According to the results of the population Rst pair wise comparisons (Supplementary Table 2), the MAP is significantly different from all other populations (3 East African Bantu, 18 West African Bantu, 2 Pygmy and 4 East African non-Bantu groups) prior to implementing the Bonferroni correction. After the application of the Bonferroni correction ($\alpha < 0.000013$), the number of significant differences between MAP and the three East Africa Bantu groups as well as between MAP and West African Bantu groups decreases to 33% and 83.3%, respectively, but all six MAP/non-Bantu comparisons remain significantly different. Overall, 5/6 (83.3%) of the uncorrected East African Bantu/East African Bantu pair wise comparisons indicate statistical significance, but this ratio is reduced to 1/6 or 16.7% (only the MAP/RWA comparison retains significance) following the Bonferroni correction.

A total of 102 (66.7%) and twelve (7.8%) of the 153 West African Bantu/West African Bantu pair wise comparisons are statistically significant before and after the Bonferroni correction, respectively. Fifty six (87.5%) of the 64 East African Bantu/West African Bantu pair wise comparisons indicate significant differences before the Bonferroni correction and this value decreases to 15 (23.4%) following its application.

Although the non-corrected results indicate that the BAK Pygmies of Gabon are very different from all other groups, including the BKO Pygmies from Cameroon (P -value = 0.0000), the Rst distances between the Pygmy collection of Cameroon and the BAK, NGU (both West African Bantus) and the BAK and TUR (both East African Bantus) become non-significant subsequent to the application of the Bonferroni correction (P -value = 0.00098). Before the Bonferroni correction, pair wise distances between the BKO Pygmies and the Bantu populations are predominantly significant (20/22 or 90.9%) (88.9% with West African Bantus and 100% with East African Bantus). Following the statistical adjustment, 45.5% of the 22 BKO/Bantu Rst distances (55.6% with West African Bantu and 50.0% with East African Bantus) remain significant.

The four non-Bantu from Tanzania (East African non-Bantu) representing Nilotic (BUR), Cushitic (DAT) and Khoisan (HAD and SAN) groups are significantly different from the remaining groups (East African Bantus, West African Bantus and Pygmies) in 95/96 (99%), and 76/96 or 79.2% of pair wise comparisons before and after the Bonferroni correction, respectively. Of the 16 East African Bantus/East African non-Bantu comparisons, 15 (93.8%) and 9 (56.3%) are significant before and after the Bonferroni correction, respectively. All six of the East African non-Bantu/East African non-Bantu pair wise comparisons indicate significant differences prior to the Bonferroni correction, but after the statistical adjustment, only three (50%), all involving the SAN, a Khoisan population, remain significant.

MDS analyses

Phylogenetic relationships among the MAP and the reference populations were assessed using MDS analysis based on Rst distances derived from the Y-STR data (Fig. 2) and confirmed by a MDS utilizing the Jaccard similarity indices (Supplementary Fig. 1). The 12-loci MDS plot (Fig. 2) displays a loose aggregation of populations mainly on the right side of the plot encompassing all of the Bantu groups including those from both East (RWA, SUK and TUR) as well as West (AKE, BEN, DUM, ESH, EVI, FAN, GAL, KOT, MAK, NDU, NGU, NZE, OBA, ORU, PUN, SHA, TEK and TSO) Africa. Although the MAP is immediately surrounded by Bantu tribes (SUK and TUR from Tanzania and ESH and MAK from Gabon), it is positioned near the upper perimeter of the Bantu cluster and close to two non-Bantu populations from Tanzania, DAT (Nilo-Saharan) and SAN (Khoisan).

The Pygmies (BAK and BKO) and two non-Bantu tribes of Tanzania (the Khoisan HAD and the Nilotic BUR), which are located to the left of the plot's origin, are isolated from the diffused Bantu cluster as well as from each other. BEN, a Bantu tribe from Gabon, is the closest group to the BKO Pygmies of Cameroon. The BAK Pygmies of Gabon, positioned to the far left of the upper left quadrant near the vertical (ordinate) origin, displays the greatest Euclidean distance away from any of the other populations. HAD and BUR occupy the upper left quadrant and are each other's nearest neighbors.

Networks and TMRCA estimates

A network representing the Bantu B2a-M150 haplogroup (Fig. 3) is a non-star like structure that, in addition to the MAP population, includes 16 West African Bantu tribes and the BKO Pygmies. It contains five multiethnic nodes, three, of which, appear near the center of the network and are composed of Bantu lineages from Gabon. The largest node, consisting of the West African Bantu groups EVI, ESH and BEN,

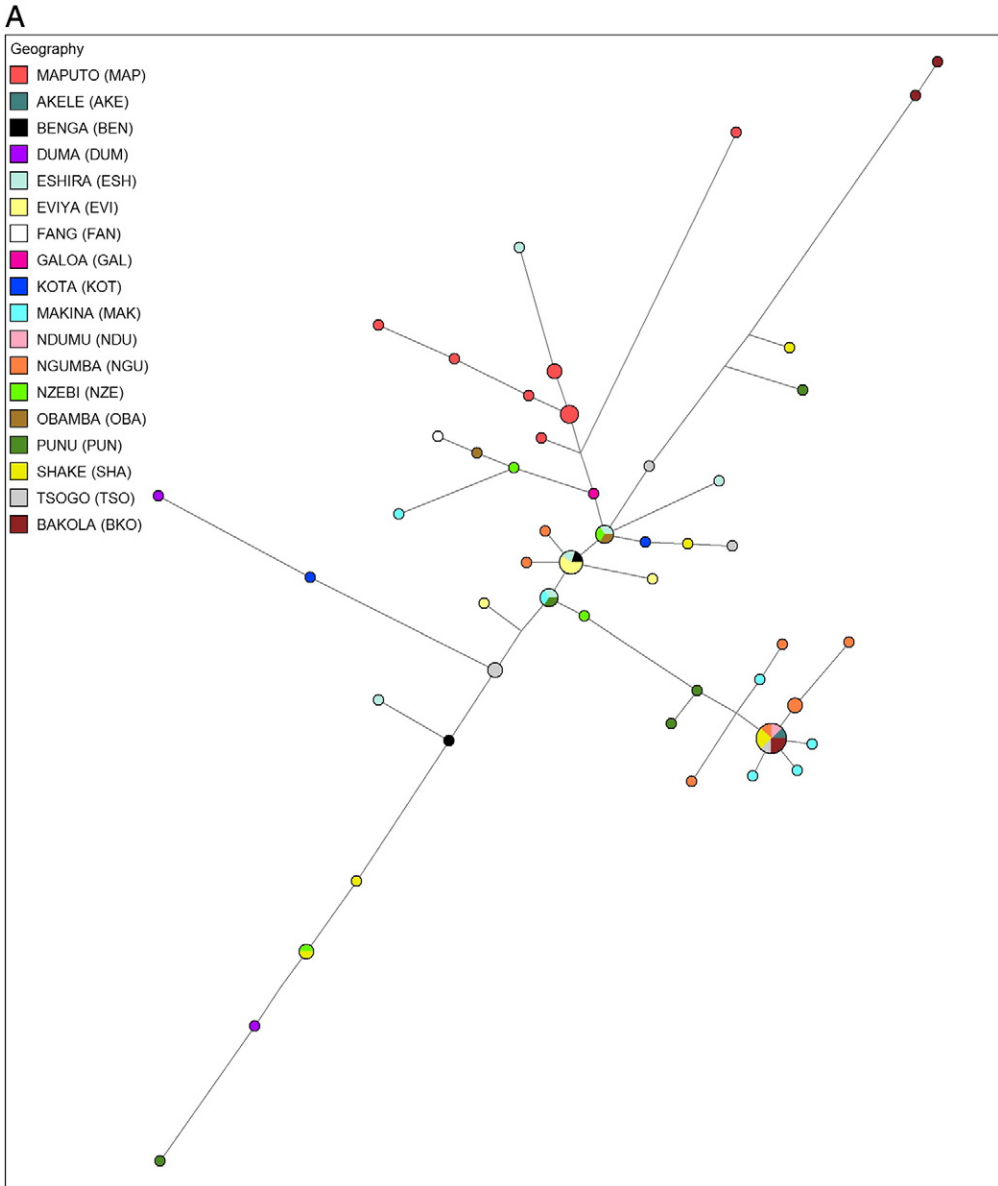


Fig. 3. (A) Network analysis based on individuals within haplogroup B2a. Colored circles represent tribal affiliations. Size of circles and length of lines are proportional to number of individuals and number of mutational steps, respectively. (B) Network analysis based on individuals within haplogroup B2b. Colored circles represent tribal affiliations. Size of circles and length of lines are proportional to number of individuals and number of mutational steps, respectively.

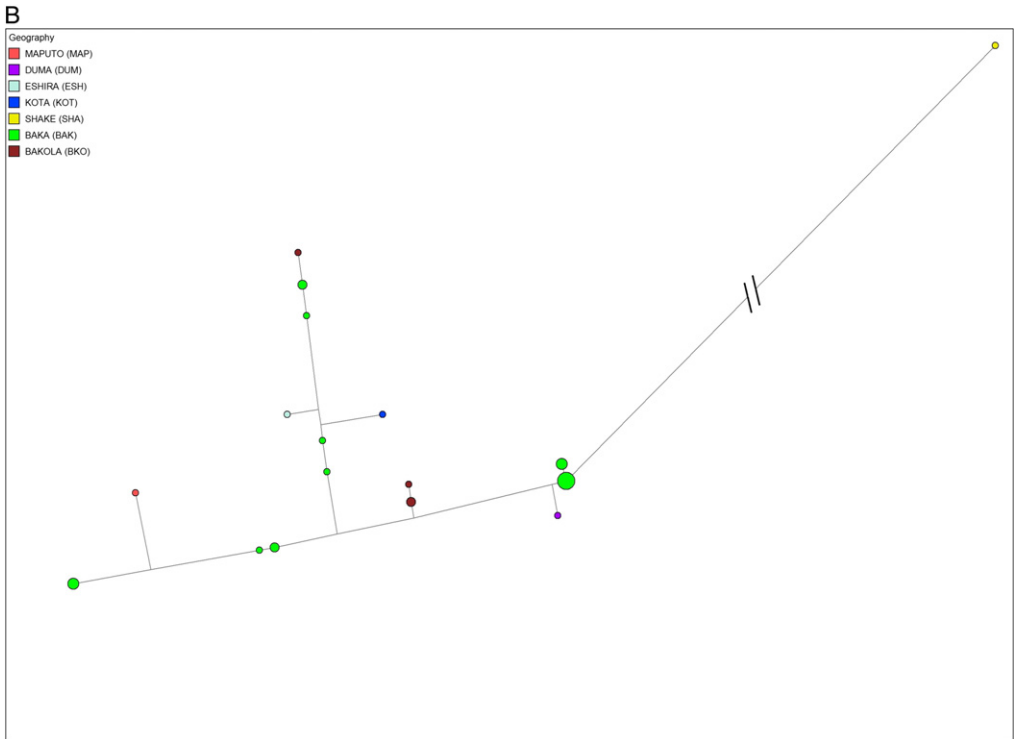


Fig. 3 (continued).

gives rise to two smaller West African Bantu centers (one node encompassing PUN, MAK and ESH and the other represented by OBA, NZE and ESH). Nearly all of the MAP B2a-M150 lineages cluster to one side of the GAL bifurcation which sprouts from the OBA/NZB/ESH node. A peripheral branch of this MAP assemblage terminates with an ESH singleton. The second offshoot of the GAL node features a NZE lineage diverging into two terminal branches (OBA and FAN singletons on one side and a MAK haplotype on the other). Other than the MAP cluster, there is little ethnic substructure. It is interesting to note that although two BKO haplotypes along the TSO branch of the OBA/NZB/ESH node are highly differentiated from the

Table 3
B2a and B2b time estimates (TE^a).

Haplogroup	N	Evol ^b TE	Gene ^c TE
<i>B2a</i>			
All	72	15,290.36 ± 2796.36	6374.86 ± 1426.97
EVI	5	3623.19 ± 1947.41	1400.00 ± 752.48
MAP	10	14,076.09 ± 7411.47	5439.00 ± 7411.47
NGU	6	7548.31 ± 3345.57	2916.67 ± 3345.57
<i>B2b</i>			
All	30	34,268.12 ± 6795.26	13,241.20 ± 2625.69
BAK	21	31,573.50 ± 5915.80	12,200.00 ± 2285.97

^a All time estimates are in years ago (ya) ± SD.

^b Dates generated from evolutionary mutation rates as in Zhivotovsky et al. (2004).

^c Dates generated from genealogical mutation rates as in Goedbloed et al. (2009).

Bantu groups, a third BKO haplotype is found along a different branch (>20 mutational steps) and is shared by five Bantu populations MAK, NDU, NGU, SHA and TSO. The evolutionary (Zhivotovsky et al., 2004) and genealogical (Goedbloed et al., 2009) TMRCA estimates for all B2a-M150 lineages are 15.3 ± 2.8 and 6.4 ± 1.4 kya, respectively, whereas those for the eleven MAP B2a-M150 haplotypes are 14.1 ± 7.4 and 5.4 ± 7.4 kya, respectively (Table 3).

The non star-like network based on the Pygmy B2b-M112 marker (Fig. 3B) encompasses fewer populations than that of the Bantu B2a-M150 haplogroup. Two Pygmy groups, BAK and BKO, supply a substantial majority (>76%) of the B2b-M112 lineages and only five Bantu populations (DUM, ESH, KOT, SHA, and MAP) are represented, each by a single haplotype. ESH and KOT, the two Bantu singletons that are closest to each other, are greater than four mutations apart and the three remaining Bantu haplotypes from different populations are widely dispersed throughout the network. The MAP population is located on a branch containing BKO lineages. The evolutionary (Zhivotovsky et al., 2004) and genealogical (Goedbloed et al., 2009) TMRCA estimates for all B2b-M112 lineages are 34.3 ± 6.8 and 13.2 ± 2.6 kya, respectively. The corresponding values for the 21 BAK haplotypes are 31.5 ± 5.9 and 12.2 ± 2.3 , respectively (Table 3).

In order to investigate the presence of the R1a1a-M198 haplogroup in the MAP population (two individuals), the genetic diversity, as reflected in the Y-STR loci, was examined utilizing network analysis (Fig. 4). The network projection based on R1a1a-M198 individuals from Eurasian populations with abundant frequencies of this haplogroup indicates considerable Y-STR diversity with minimal haplotype sharing. The vast majority of the populations are represented by singleton haplotypes separated by 1 to 10 mutational steps. The two samples from the MAP population exhibit identical haplotypes and are 4 mutational steps removed from the nearest South Asian haplotype. The relative similarity of the MAP haplotype to the rest of the haplotypes from Eurasian populations suggests that the presence of R1a1a-M198 in MAP is the result of recent migration. Historically, East Africa has been part of the dominions of a number of empires including the Kingdom of Oman and intense commerce, including the East African Slave Trade, has characterized the two regions. It is likely that these two individuals with identical haplotypes arrived in Maputo in recent times and are not the result of ancient back migration.

AMOVA analyses

AMOVAs were performed on the MAP and a total of 27 geographically relevant reference populations. In the geographical AMOVA (Table 4), the populations were partitioned into three geographical regions (West Africa, East Africa and Southeast Africa). In this analysis significant correlations were found with YSTR variation on all three levels, among geographical groups (F_{CT}), among populations within geographical groups (F_{SC}) and within populations (F_{ST}), all at P -value < 0.00001. Likewise, the results of the language based AMOVA (Table 4) also generated significant correlations with YSTR variation among the five language groups (F_{CT}) at P -value = 0.01271, among the populations within each group (F_{SC}) at P -value < 0.00001 as well as within populations (F_{ST}) at P -value < 0.00001. Variation among the sub-Saharan language groups is lower than that exhibited by tribes within groups.

Discussion

In the present study, 78 Bantu males from the Maputo Province of Mozambique (MAP), a region believed to be part of the southeastern fringe of the Bantu expansion, were genotyped utilizing 12 Y-STR loci (DYS19, DYS385a/b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, and DYS439) against the backdrop of the binary Y-SNP marker system. Along with Bantu-speaking MAP, the phylogenetic analyses include 27 geographically-targeted reference populations in order to ascertain the genetic relationships between the MAP population and other sub-equatorial tribes from West and East Africa (Bantu, Pygmy, Khoisan, Nilotic and Cushitic) (Table 1). Our objective is twofold. First, we aim to assess the relative genetic contributions by Bantu versus ancient non-Bantu (e. g., Pygmy, Khoisan, Nilotic and Cushitic) people to the MAP Bantu and second, to examine genetic affinities among the reference populations in order to shed light on the possible routes of the Bantu expansion.

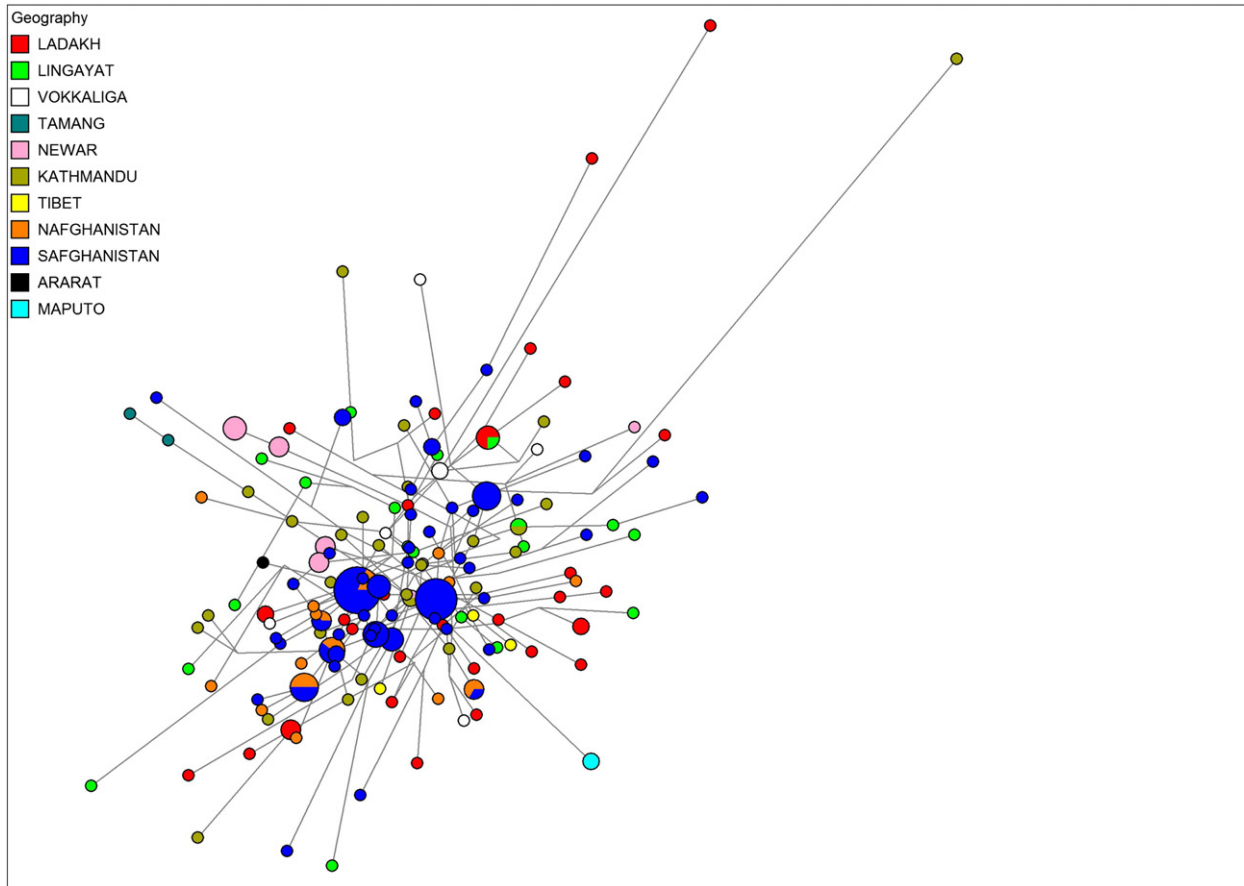


Fig. 4. Network based on R1a1a-M198 haplogroup. References of populations: Ladakh (Chennakrishnaiah et al., in preparation), Lingayat (Chennakrishnaiah et al., 2013), Vokkaliga (Chennakrishnaiah et al., 2013), Tamang (Gayden et al., 2007), Newar (Gayden et al., 2007), Kathmandu (Gayden et al., 2007), Tibet (Gayden et al., 2007), Nafghanistan (Lacau et al., 2011), Safghanistan (Lacau et al., 2011), Ararat (Herrera et al., 2011). Colored circles represent tribal affiliations. Size of circles and length of lines are proportional to number of individuals and number of mutational steps, respectively.

Table 4
Analysis of molecular variance (AMOVA) using Y-STRs.

Group	Number of groups	Within populations		Among populations within groups		Among groups	
		Variation (%)	Φ_{ST}	Variation (%)	Φ_{SC}	Variation (%)	Φ_{CT}
Language ^a	5	73.34	0.27*	5.09	0.06*	21.57	0.22**
Geography ^{a,b}	3	84.15	0.16*	13.83	0.14*	2.03	0.02*

* P -value < 0.00001.

** P -value = 0.01271.

^a Linguistic partitioning (5 groups): Nilotic, Cushitic, Khoisan, Pygmy, Bantu.

^b Geographical partitioning (3 groups): West Bantu, East Bantu, Southeast Bantu.

East African Bantu and West African Bantu

With respect to paternal lineages, the East African Bantu and West African Bantu groups are more similar to each other than to the non-Bantu tribes (Cushitic, Khoisan, Nilotic and Pygmy). This genetic affinity can be seen in the MDS analysis (Fig. 2) in which the East and West Bantu collections form a large singular assembly occupying the right area of the plot. Furthermore, this similarity among the Bantu groups is apparent in the relative percentages of significant different pair wise comparisons (23.4% versus percentages ranging from 50% to 100% for the eight Bantu/non-Bantu comparison sets). The Bantu expansion marker, E1b1a-M2, and the major component of the Bantu Y-SNP frequency distributions (Berniell-Lee et al, 2009; Tishkoff et al, 2007), coupled with relatively low levels of ancient lineages (A and B2b), are the common denominators among all the Bantu populations examined.

Genetic affinities of MAP to East African and West African Bantus, and non-Bantus

The MAP population is composed primarily of Bantu lineages (95%) including a high percentage of the Bantu expansion signature markers, E1b1a-M2 (71.8%) and B2a-M150 (14.1%) (Table 2) and exhibits a relatively close paternal genetic affinity with both East African and West African Bantu tribes. These genetic similarities are reflected both, by the Euclidean position of the MAP (surrounded by both East and West African Bantu populations) in the MDS plot (Fig. 2), and in the relative low percentages of statistical significant differences in the Rst population pair wise comparisons (Supplementary Table 2) (33.3% and 83.3% after Bonferroni correction for East and West African Bantu groups, respectively) versus 100% of MAP/non-Bantu comparisons. In addition to the preeminent Bantu genetic heritage, the haplogroup frequency distribution (Table 2) also reveals that only traces of Khoisan (1.3%) and Pygmy (1.3%) markers persist in the MAP gene pool. Further, the highly derived E1b1b1-M35 haplogroup, observed at polymorphic frequencies in all Tanzanian tribes examined including the East African Bantu (6% in Sukuma and 15% in Turu) and in the four East African non-Bantu groups (33% in Burunge, 54% in Datog, 16% in Hadza 16% and 34% in Sandawe) is also detected in the MAP (1.3%). These observations suggest that, in addition to affiliation due to a shared language, cultural heritage and genetic origin, geography (e.g., east versus west) has impacted the paternal landscape of East and Southeast Africa. This data is echoed by the AMOVA results that indicate statistical correlations between genetic diversity and geography as well as language. The fact that the variation among the sub-Saharan language groups is lower than that exhibited by populations within groups may reflect the strong genetic impact of the Bantu expansion on non-Bantu-speaking populations as well as the high language diversity in specific geographical regions.

The asymmetrical abundance of the E1b1b1-M35 haplogroup in the Tanzanian versus the MAP males (a mean of 26% in the six Tanzanian populations in contrast to 1.3% in the MAP) along with its total absence in 17 out of 18 West African Bantu tribes (2.1% in Galoa) and the two Pygmy tribes may hint at a north to south trajectory along the eastern leg of the Bantu expansion. This data is congruent with an independent migration from North Cameroon, following a southeastern trajectory to Southeast Africa in addition to the southwestern course instead of only an initial single southwestern migration from the Bantu homeland to Southwest Africa and a subsequent, more recent, longitudinal dispersion eastward. However, since alternative explanations are feasible, such as one or multiple episodes of post-expansion

interactions between the MAP and their northern Tanzanian neighbors, more data is needed to determine the significance of this shared haplogroup.

The Bantu haplogroup marker B2a-M150 is detected at abundant levels (14.1%) in the MAP population. The network analysis of individuals in this haplogroup exhibits an absolute partitioning of B2a-M150 MAP males to a single branch of the projection (Fig. 3), although the rest of the populations are randomly distributed. This segregation pattern suggests limited gene flow into the Bantus from the Maputo Province from other West and East Bantu populations. This clustering of MAP individuals reflecting relative genetic homogeneity is compatible with a limited number of migrations reaching the Southeast fringes of its expansion.

Also noteworthy is the trace of R1a1a-M198 (2.6%), a member of the Eurasian R1a-M420 branch in the MAP population. R1b1a-P297, a related Eurasian marker (R1a1a-M198 and R1b1a-P297 are members of the Eurasian haplogroup R1 characterized by the presence of the M173 polymorphism (Underhill et al., 2000)), also detected in the west African Bantu groups (ranging from 2% to 20%), in the two Pygmy groups (3% and 4.5% in Baka and Bakola, respectively) and possibly in the Nilotic and Cushitic (8% and 6% of the C, F–R haplogroups in Burunge and Datog, respectively) (Berniell-Lee et al., 2006, 2009), may represent Eurasian back migration events (Cruciani et al., 2002) or more recent gene flow from Eurasian settlers.

Conclusion

According to the haplogroup and haplotype frequency distributions as well as the MDS and the pair wise Rst comparison analyses, the Maputo population, a representative of the extreme southeastern fringe of the Bantu expansion, displays a close paternal affinity to both eastern and western Bantu populations due to its high proportion of Bantu Y chromosomal markers. Only traces of Khoisan (1.3%) and Pygmy (1.3%) markers persist in the Maputo Bantu gene pool. The shared presence of haplogroup E1b1b1-M35 in all Tanzanian tribes examined, including Bantu and non-Bantu groups, in conjunction with its nearly complete absence in the West African populations indicate that, in addition to a shared linguistic, cultural and genetic heritage, geography (east vs. west) may have impacted the paternal genetic landscape of sub-Saharan Africa. Also, the occurrence of R1a1a-M17/M198, a member of the Eurasian R1a-M420 branch in the Maputo population, likely represent recent admixture events since R1a-M420 is present in both Europe and India and the two R1a1a MAP individuals exhibit close YSTR similarity to the Eurasian populations examined. Overall, the admixture and assimilation processes of Bantu elements into native populations have been both highly complex and region-specific. Two theories have been proposed to explain the dispersal of the Bantus in their route to populate Southeast Africa. In one, two independent demographic expansions took place. One proceeded along a southwestern course emanating from their homeland in West Africa. A second migration wave, also from North Cameroon, followed a southeastern trajectory reaching the fringes of the expansion in Southeast Africa as recent as 300 ya. A second proposal argues for an initial single southwestern migration from the Bantu homeland and a subsequent longitudinal dispersion eastward.

The presence of the E1b1b1-M35 haplogroup in the Tanzanian and MAP males along with its total absence in 17 out of 18 West African Bantu tribes and two Pygmy tribes from West Africa suggest a north to south trajectory along a putative eastern leg of the Bantu expansion. This data is congruent with an independent migration from the Bantu homeland following a southeastern trajectory to Southeast Africa in addition to the southwestern course instead of only an initial single southwestern migration from the Bantu homeland to Southwest Africa and a subsequent, more recent, longitudinal dispersion eastward. However, since alternative explanations are feasible, such as one or multiple episodes of post-expansion interactions between the MAP and their northern Tanzanian neighbors, more data is required to assess the significance of this shared haplogroup.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.mgene.2014.08.003>.

References

- Alves, C., Gusmão, L., Barbosa, J., Amorim, A., 2003. Evaluating the informative power of Y-STRs: a comparative study using European and new African haplotype data. *Forensic Sci. Int.* 134, 126–133.
- Alves, I., Coelho, M., Gignoux, C., Damasceno, A., Prista, A., Rocha, J., 2011. Genetic homogeneity across Bantu-speaking groups from Mozambique and Angola challenges early split scenarios between East and West Bantu populations. *Hum. Biol.* 13–38.

- Balamurugan, K., Duncan, G., 2012. Y chromosome STR allelic and haplotype diversity in a Rwanda population from East Central Africa. *Leg. Med. Tokyo* 14, 105–109.
- Batai, K., Babrowski, K.B., Arroyo, J.P., Chapurukha, C.M., Kusimba, C., Williams, S.R., 2013. Mitochondrial DNA diversity in two ethnic groups in Southeastern Kenya: perspectives from the northeastern periphery of the Bantu expansion. *Am. J. Phys. Anthropol.* 150, 482–491.
- Batini, C., Coia, V., Battaglia, C., Rocha, J., Pilkington, M.M., Spedini, G., Comas, D., Destro-Bisol, G., Calafell, F., 2007. Phylogeography of the human mitochondrial L1c haplogroup: genetic signatures of the prehistory of Central Africa. *Mol. Phylogenet. Evol.* 43, 635–644.
- Beleza, S., Gusmão, L., Amorim, A., Carracedo, A., Salas, A., 2005. The genetic legacy of western Bantu migrations. *Hum. Genet.* 117, 366–375.
- Berniell-Lee, G., Bosch, E., Bertranpetit, J., Comas, D., 2006. Y-chromosome diversity in Bantu and Pygmy populations from Central Africa. *Int. Congr. Ser.* 1288, 234–236.
- Berniell-Lee, G., Calafell, F., Bosch, E., Heyer, E., Sica, L., Mouguiama-Daouda, P., van der Veen, L., Hombert, J., Quintana-Murci, L., Comas, D., 2009. Genetic and demographic implications of the Bantu Expansion: insights from human paternal lineages. *Mol. Biol. Evol.* 26, 1581–1589.
- Calderon, S., Perez-Benedico, D., Mesa, L., Guyton, D., Rowold, D., Herrera, R.J., 2013. Phylogenetic and forensic studies of the Southeast Florida Hispanic population using the next-generation forensic PowerPlex® Y23 STR marker system. *Leg. Med. Tokyo* 15, 289–292.
- Carvalho, M., Brito, P., Lopes, V., Andrade, L., Anjos, M.J., Corte Real, F., Gusmão, L., 2010. Analysis of paternal lineages in Brazilian and African populations. *Genet. Mol. Biol.* 33, 422–427.
- Cavalli-Sforza, L.L., 1986. *African Pygmies*. Academic Press, Orlando, FL.
- Chennakrishnaiah, S., Perez, D., Gayden, T., Rivera, L., Regueiro, M., Herrera, R.J., 2013. Indigenous and foreign Y-chromosomes characterize the Lingayat and Vokkaliga populations of Southwest India. *Gene* 526, 96–106.
- Chennakrishnaiah, S., Rowold, D.J., Perez, D., Gayden, T., Herrera, R.J., 2014n. On the Genetics of the Silk Road. (in preparation).
- Cruciani, F., Santolamazza, P., Shen, P., Macaulay, V., Moral, P., Olckers, A., Modiano, D., Holmes, S., Destro-Bisol, G., Coia, V., Wallace, D., Oefner, P., Torroni, A., Cavalli-Sforza, L.L., Scozzari, R., Underhill, P., 2002. A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *Am. J. Hum. Genet.* 70, 1197–1214.
- De Filippo, C., Koen, B., Stoneking, M., Pakendorf, B., 2012. Bringing together linguistic and genetic evidence to test the Bantu expansion. *Proc. Roy. Soc. Lond. B* 279, 3256–3263.
- Desmond, J., Brandt, S., 1984. *From Hunters to Farmers: The Causes and Consequences of Food Production in Africa*. University of California Press.
- Diamond, J., 1997. *Guns, Germs, and Steel: The Fates of Human Societies*. W.W. Norton & Company, New York.
- Ehret, C., 2001. Bantu expansions: re-envisaging a central problem of early African history. *Int. J. Afr. Hist. Stud.* 34, 5–41.
- Excoffier, L., Laval, L., Schneider, S., 2005. Arlequin ver.3.0: an integrated software package for population genetic data analyses. *Evol. Bioinforma.* 1, 47–50.
- Fluxus Engineering, 2012. NETWORK 4.6.1.1 software program. <http://www.fluxus-engineering.com>.
- Gayden, T., Cadenas, A., Regueiro, M., Singh, N., Zhivotovskiy, L., Underhill, P., Cavalli-Sforza, L.L., Herrera, R.J., 2007. The Himalayas as a directional barrier to gene flow. *Am. J. Hum. Genet.* 80, 884–894.
- Goedbloed, M., Vermeulen, M., Fang, R., Lembring, M., Wollstein, A., Ballantyne, K., Lao, O., Brauer, S., Krüger, C., Roewer, L., Lessig, R., Ploski, R., Dobosz, T., Henke, L., Henke, J., Furtado, M., Kayser, M., 2009. Comprehensive mutation analysis of 17 Y chromosomal short tandem repeat polymorphisms included in the AmpFISTR Yfiler PCR amplification kit. *Int. J. Legal Med.* 123, 471–482.
- Greenberg, J., 1972. Linguistic evidence regarding Bantu origins. *J. Afr. Hist.* 13, 189–216.
- Gusmão, L., Sanchez-Diz, P., Alves, C., et al., 2003. Results of the GEP-ISFG collaborative study on the Y chromosome STRs GATA A10, GATA C4, GATA H4, DYS437, DYS438, DYS439, DYS460 and DYS461: population data. *Forensic Sci. Int.* 135, 150–157.
- Hammer, M., Horai, S., 1995. Y chromosomal DNA variation and the peopling of Japan. *Am. J. Hum. Genet.* 56, 951–962.
- Herrera, R.J., Lowery, R.K., Hadden, L., Silvia Calderon, S., Chiou, C., Yepiskoposyan, L., Regueiro, M., Underhill, P.A., Herrera, R.J., 2011. Neolithic patrilineal signals indicate that the Armenian plateau was repopulated by agriculturalists. *Eur. J. Hum. Genet.* 20, 313–320.
- Karafet, T., Mendez, F., Meilerman, M., Underhill, P., Zegura, S., Hammer, M., 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res.* 18, 830–838.
- Kayser, M., Brauer, S., Schädlich, H., Prinz, M., Batzer, M., Zimmerman, P., Boatman, B., Stoneking, M., 2003. Y chromosome STR haplotypes and the genetic structure of U.S. populations of African, European, and Hispanic ancestry. *Genome Res.* 13, 624–634.
- Lacau, H., Bukhari, A., Gayden, T., La Salvia, J., Regueiro, M., Stojkovic, O., Herrera, R.J., 2011. Y-STR profiling in two Afghanistan populations. *Legal Med.* 13, 103–108.
- Luis, J., Rowold, D., Regueiro, M., Caeiro, B., Cinnioglu, C., Roseman, C., Underhill, P., Cavalli-Sforza, L.L., Herrera, R.J., 2004. The Levant versus the Horn of Africa: evidence for bidirectional corridors of human migrations. *Am. J. Hum. Genet.* 74, 532–544.
- Martinez, L., Underhill, P.A., Zhivotovskiy, L.A., Conti, S., Elisabetta Mamolini, E., Cavalli-Sforza, L.L., Herrera, R.J., 2007. Paleolithic Y-haplogroup heritage predominates in a Cretan highland plateau. *Eur. J. Hum. Genet.* 15, 485–493.
- Myres, N., Rootsi, S., Lin, A., Järve, M., King, R., Kutuev, I., Cabrera, V., Khusnutdinova, E., Pshenichnov, A., Yunusbayev, B., Balanovsky, O., Balanovska, E., Rudan, P., Baldovic, M., Herrera, R.J., Chiaroni, J., Di Cristofaro, J., Vilems, R., Kivisild, T., Underhill, P., 2011. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *Eur. J. Hum. Genet.* 19, 95–101.
- Newman, J., 1995. *The Peopling of Africa: A Geographic Interpretation*. Yale University Press, New Haven, CT.
- Oliver, R., 2009. The problem of the Bantu expansion. *J. Afr. Hist.* 7, 361.
- Pakendorf, B., Bostoen, K., de Filippo, C., 2011. Molecular perspectives on the Bantu expansion: a synthesis. *Lang. Dyn. Chang.* 1, 50–88.
- Pereira, L., Gusmão, L., Alves, A., Amorim, A., Prata, M.J., 2002. Bantu and European Y-lineages in Sub-Saharan Africa. *Ann. Hum. Genet.* 66, 369–378.
- Phillipson, D., 1993. *African Archaeology*. Cambridge University Press, Cambridge MA.
- Quintana-Murci, L., Quach, H., Harmant, C., et al., 2008. Maternal traces of deep common ancestry and asymmetric gene flow between Pygmy hunter-gatherers and Bantu-speaking farmers. *PNAS* 105, 1596–1601.
- Regueiro, M., Cadenas, A., Gayden, T., Underhill, P., Herrera, R.J., 2006. Iran: tricontinental nexus for Y-chromosome driven migration. *Hum. Hered.* 61, 132–143.

- Regueiro, M., Alvarez, J., Rowold, D., Herrera, R.J., 2013. On the origins, rapid expansion and genetic diversity of Native Americans from hunting-gatherers to agriculturalists. *Am. J. Phys. Anthropol.* 150, 333–348.
- Rowold, D.J., Herrera, R.J., 2003. Inferring recent human phylogenies using forensic STR technology. *Forensic Sci. Int.* 133, 260–265.
- Russell, T., Silva, F., Steele, J., 2014. Modelling the spread of farming in the Bantu-speaking regions of Africa: an archaeology-based phylogeography. *PLoS One* 9, e87854.
- Sanchez-Diz, P., De La Fe, T., Quintans, B., Salas, A., Lareu, M., Carracedo, A., 2003. Y-chromosome STRs in populations of Bantu origin from Mozambique: male contribution to the Africa genetic pool and forensic implications. *Int. Congr. Ser.* 1239, 419–424.
- Tishkoff, S., Gonder, M., Henn, B., Mortensen, H., Knight, A., Cignoux, C., Fernandopulle, N., Lema, G., Nyambo, T., Ramakrishnan, U., Reed, F., Mountain, J., 2007. History of click-speaking populations of Africa inferred from mtDNA and Y chromosome genetic variation. *Mol. Biol. Evol.* 24, 2180–2195.
- Underhill, P., Shen, P., Lin, A., Jin, L., Passarino, G., Yang, W., Kauffman, E., Bonn -Tamir, B., Bertranpetit, J., Francalacci, P., Ibrahim, M., Jenkins, T., Kidd, J., Mehdi, S., Seielstad, M., Wells, R., Piazza, A., Davis, R., Feldman, M., Cavalli-Sforza, L.L., Oefner, P., 2000. Y chromosome sequence variation and the history of human populations. *Nat. Genet.* 26, 358–361.
- Underhill, P., Myres, N., Rootsi, S., Metspalu, M., Zhivotovsky, L., King, R., Lin, A., Chow, C., Semino, O., Battaglia, V., Kutuev, I., J rve, M., Chaubey, G., Ayub, Q., Mohyuddin, A., Mehdi, S., Sengupta, S., Rogaev, E., Khusnutdinova, E., Pshenichnov, A., Balanovsky, O., Balanovska, E., Jeran, N., Augustin, D., Balovic, M., Herrera, R.J., Thangaraj, K., Singh, V., Singh, L., Majumder, P., Rudan, P., Primorac, D., Villems, R., Kivisild, T., 2010. Separating the post-glacial coancestry of European and Asian Y chromosomes within haplogroup R1a. *Eur. J. Hum. Genet.* 18, 479–484.
- Vansina, J., 1995. New linguistic evidence and the Bantu Expansion. *J. Afr. Hist.* 36, 173–195.
- Y Chromosome Consortium., 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res.* 12, 339–348.
- Zhivotovsky, L., Underhill, P., Cinnioglu, C., Kayser, M., Morar, B., Kivisild, T., Scozzari, R., Cruciani, F., Destro-Bisol, G., Spedini, G., Chambers, G.K., Herrera, R.J., Yong, K., Gresham, D., Tournev, I., Feldman, M., Kalaydjieva, L., 2004. The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. *Am. J. Hum. Genet.* 74, 50–61.